

Sigla Asignatura	BDY710 1	Nombre de la Asignatura	BIG DATA
Nombre del Recurso Didáctico	1.1.8_ACT_Interacción con Hadoop Distributed file System.		
Unidad de Aprendizaje N° 1	Ecosistema Hadoop		
Unidades de Competencia	Administra el volumen, la diversidad y complejidad del almacenamiento de datos para gestionar y extraer valor de acuerdo a los requerimientos de la organización.		

Actividad 1.1.8: Interacción con Hadoop Distributed File System

Descripción

En esta actividad, ejercitará el uso del sistema de archivos de Hadoop, HDFS, mediante la máquina virtual montada en sus computadores. Utilizando los comandos descritos en la siguiente figura y a través de una terminal, deberán interactuar con distintas funcionalidades de HDFS. El modo de uso de estos comandos es el siguiente: `hdfs dfs comando`, donde comando es uno de los siguientes:

Command	Description
<code>-ls path</code>	Lists contents of directory
<code>-lsr path</code>	Recursive display of contents
<code>-du path</code>	Shows disk usage in bytes
<code>-dus path</code>	Summary of disk usage
<code>-mv src dest</code>	Move files or directories within HDFS
<code>-cp src dest</code>	Copy files or directories within HDFS
<code>-rm path</code>	Removes the file or empty directory in HDFS
<code>-rmr path</code>	Recursively removes file or directory
<code>-put localSrc dest</code> (Also <code>-copyFromLocal</code>)	Copy file from local filesystem into HDFS

Command	Description
<code>-get src localDest</code>	Copy from HDFS to local filesystem
<code>-cat filename</code>	Display contents of HDFS file
<code>-tail file</code>	Shows the last 1KB of HDFS file on stdout
<code>-chmod [-R]</code>	Change file permissions in HDFS
<code>-chown [-R]</code>	Change ownership in HDFS
<code>-help</code>	Returns usage info

Instrucciones

1. Crear un directorio en HDFS, dentro del home del usuario actual.
2. Listar el contenido de un directorio en HDFS que no se encuentre vacío.
3. Copiar un archivo, desde el sistema de archivos local, a HDFS.
4. Mover y copiar archivos dentro de HDFS.
5. Verificar información sobre el uso del disco por parte de HDFS.
6. Borrar un archivo y un directorio en HDFS.

Además, con el fin de verificar los efectos sobre el sistema de archivos, puede utilizar los comandos `hdfs fsck filename` y `hdfs dfsadmin report`, para revisar el estado de los archivos copiados y del sistema, respectivamente.

Recursos

- Máquina virtual de Cloudera (sandbox).

Criterio de evaluación

1. Ejecuta los comandos correctamente. (5 puntos)
2. Documenta los comandos usados y los resultados obtenidos. (5 puntos)